SYNTHETIC REALITIES: AI-GENERATED DEEPFAKES AND CONSPIRACY THEORIES AS A CHALLENGE TO TRUST IN MODERN DEMOCRACIES

Kornelia BATKO[1]

This chapter explores the complex interplay between artificial intelligence (AI), deepfake technology, and the spread of conspiracy theories, highlighting their collective impact on democratic trust. The analysis conducted in the chapter on using synthetic reality (hyperrealistic images, audio and video files generated by artificial intelligence) to spread disinformation and conspiracy narratives has made it possible to identify threats and consequences for democratic societies. Based on the referenced cases of deepfake use in political and social contexts, the chapter discusses how these technology-based manipulations support the spread of conspiracy theories, thereby exacerbating social tensions, undermining public trust in democratic institutions, and disrupting political discourse. The chapter emphasises the need to develop proactive strategies to limit the spread of AI-powered conspiracy theories to protect trust in democracy and social resilience.

Key words: conspiracy theories; AI-generated content; synthetic realities; deepfakes; trust.

1 Kornelia Batko, associate professor at the Institute of Political Science, Faculty of Social Sciences, University of Silesia, Katowice, Poland. This research was funded by the National Science Centre, Poland (grant no. 2020/39/I/HS5/00176).

Introduction

We live in an era where information is a valuable resource and tool for influencing others. Information is the foundation of democratic societies, and the line between true and false is becoming increasingly blurred. The development of artificial intelligence (AI) in recent years has affected many sectors of the economy, as well as our daily lives, bringing enormous benefits and improving the quality of life. However, with the evolution of AI, unprecedented challenges are also emerging. One of these challenges is synthetic reality, applications of artificial intelligence, with AI-generated content (Wang, 2023). Synthetic Realities (SR) can be defined as computer-generated and hyper realistic media content, especially deepfakes (George and George, 2023; Hynek et al., 2025). Deepfakes are audiovisual materials (images, audio and video recordings) produced using advanced generative artificial intelligence to imitate real people or events (Babaei et al., 2025). The fast development of machine learning has made deepfakes increasingly realistic and challenging to distinguish from reality (Sharma et al., 2025). In this way, a digital equivalent of a person can be created and used to produce fake multimedia content, including by placing digitally generated people in events that did not take place, or imitating the events themselves in a way that is difficult to distinguish from the authentic version. That kind of AI-generated content is created not only for entertainment but more often for shaping public opinion, influencing election results, or destabilising democratic institutions. Citizens can no longer reliably determine whether the information they encounter is true, which has profound implications for public trust. Thus, this technology spreads political misinformation, propaganda, and conspiracy theories. Synthetic content in the form of deepfakes, which reinforces conspiracy theories and undermines public trust, is a significant threat to the performance of modern democracies.

Conspiracy theories (CTs) are characterised as explanatory narratives regarding powerful entities conspiring covertly to achieve malevolent goals (Fortaleza, 2020; Motta, 2021). Although significant organisations could be accused of conspiracies the government and multinational businesses remain the most commonly implicated (Douglas et al., 2019). Accordingly, 'conspiracy beliefs' describes convictions regarding specific conspiracy theories (Pilch et al., 2023). Conspiracy theories have been part of political and social life for a long time. Historically, they have emerged during periods of uncertainty, serving as narratives through which individuals have attempted to explain complex events or

crises. While some conspiracy theories may be harmless or amusing, spreading harmful ones has been linked to political violence, public crises, and a decline in trust in democratic governance. In today's world, thanks to content generated by artificial intelligence, which is an effective medium for such narratives, they are becoming more convincing and difficult to refute.

The interaction between deepfakes and conspiracy theories is a huge problem, as deepfakes provide visual and audio 'evidence' to make conspiracy theories seem credible, even if they are fake. This chapter analyses the interaction between synthetic realities (deepfake technology), the spread of conspiracy theories, and the resulting challenges to trust in democracies. Examples of incidents involving deepfakes in various political contexts – from global conflicts, through Western countries (Europe and the US), to Poland – will be discussed to show how synthetic media can intensify social tensions, reinforce misinformation, and fuel conspiracy theories. To understand these relationships, two research questions were formulated:

- How do deepfakes increase the spread and impact of conspiracy theories in democratic societies?
- How do synthetic realities affect public trust and disrupt political discourse?

Based on a series of case studies and theoretical perspectives from political science, psychology, and information technology, this chapter argues that addressing the challenge of deepfakes and conspiracy theories requires a series of coordinated actions. Thus, the spread of conspiracy theories aided by synthetic realities risks creating a post-truth environment in which trust will be irreversibly damaged, and democracy itself may be endangered. This is particularly relevant today, as we face multiple crises: the COVID-19 pandemic and its consequences, the war in Ukraine, the Israeli-Palestinian conflict, and a deepening migration crisis.

Theoretical Background

Conspiracy Theories

The phenomenon of conspiracy theories has been a part of human societies for centuries, explaining complex events by hypothesising the existence of covert, malevolent actions by powerful actors. A conspiracy theory can be defined as

'a secret arrangement between a small group of actors to usurp political or economic power, violate established rights, hide vital secrets, or illicitly cause widespread harm' (Boncu et al., 2022). According to Swami and Furnham (2014) 'conspiracy theory' usually refers to *a subset of false beliefs in which the ultimate cause of an event is believed to be due to a plot by multiple actors working together with a clear goal in mind, often unlawfully and in secret*. Uscinski et al. (2016) define conspiracy belief as *'an individual's belief in a specific conspiracy theory'.* Furthermore, they state that 'conspiratorial predispositions or thinking' can be defined as 'an individual's underlying propensity to view the world in conspiratorial terms'.

The concept of conspiracy theories frequently ascribes extraordinary capabilities to specific entities, such as planning, controlling others, and keeping secrets (Sunstein and Vermeule, 2009). The three elements that are typically present in conspiracy theories are: Actors (usually said to be powerful elites, institutions, or secret organisations), Intentionality (the belief that actions are deliberate and hidden) and Secrecy (the assumption that the true motives are concealed from the public).

Van Prooijen and Douglas (2017) extend this definition by viewing conspiracy theories as cognitive narratives that help individuals make sense of a situation. It follows that such beliefs are increasingly likely to the extent that people experience uncertainty or a lack of control. This explains the spread of conspiracy theories in times of crisis in society, as such situations are perceived as uncontrollable and therefore cause considerable uncertainty and anxiety among citizens.

In the digital age, conspiracy theories have transformed into global narratives that spread rapidly via online platforms, creating a cascade effect. The democratisation of information production, combined with the rapid spread of digital communication, has meant that even marginal conspiracy theories can attract public attention within hours. Sunstein and Vermeule (2009) linked the cascade effect to group polarisation, which characterises a situation in which members of a decision-making group tend to adopt a more radical position in line with their pre-decision views. Thus, belief in conspiracy theories is often fuelled by group polarisation.

Social media platforms such as Facebook, Twitter, TikTok, and YouTube act as accelerators for conspiracy narratives, allowing the viral spread of emotionally charged content, rewarding user engagement through likes, shares, and subscriptions, which encourages sensationalism and misinformation, and removing

traditional gatekeepers (e.g., journalists, editors, moderators) who have historically filtered information. Due to all this, fake news spreads faster and more broadly than fact-based information, mainly because it elicits stronger emotional responses (Wrzosek, 2019). Conspiracy theories fit well into the digital environment because they often involve dramatic revelations, secret plots, and big money. Social media algorithms also support the reinforcement of fake content as they are designed to maximise engagement and inadvertently create filter bubbles and echo chambers, where users are repeatedly exposed to duplicate content and narratives. This further exacerbates polarisation and makes conspiracy beliefs more resistant to refutation. Moreover, social media has fundamentally changed how conspiracy theories are created and spread.

Synthetic Realities and Deepfakes

The definition of synthetic realities evolved from a narrow technological understanding to a broader socio-cultural concept focused on human relationships with AI-generated content. Some definitions emphasise the technological nature of synthetic realities. Lusquino Filho and Rocha define them as digital constructs or enhancements created using artificial intelligence techniques that use deep learning and data analysis algorithms to build a new 'reality' or narrative, regardless of whether they may mislead the audience (Filho and Rocha, n.d.). Other researchers noted the immersive nature of synthetic realities and their impact on user perception. Cardenuto et al. (2023) emphasise that synthetic realities have become a 'new wave' of human-technology interaction, in which the boundary between reality and artificiality becomes quite fluid. Subsequent approaches emphasise their application and the ambivalent potential of these tools – both creative and disinformative. Thus, Wang (2023) defines synthetic realities as products of the convergence of artificial intelligence, immersive media, and automation, which enable the creation of realistic virtual experiences, but at the same time open up space for manipulation, disinformation, and privacy violations. Finally, Flores (2025) added an ethical and communicative dimension to this concept, pointing out that synthetic realities redefine the concepts of authenticity, consent, and trust, becoming one of the key challenges for contemporary digital democracy. All kinds of hyper-realistic AI-generated content blur the boundaries between reality and fake, redefining the concepts of consent, authenticity, and trust in communication (Stroebel et al., 2023)

Unlike traditional images, films, or computer-generated images, which require considerable expertise and resources, synthetic reality is created using machine learning models capable of independently generating realistic images, sounds, and texts. A characteristic feature of synthetic reality is that it is realistic, which makes it so convincing that the recipient perceives it as authentic. For these reasons, synthetic reality can be seen as a potent persuasion, manipulation, and deception tool. The development of synthetic media is based on deep learning technologies, including neural networks such as GANs (Generative Adversarial Networks), VAEs (Variational Autoencoders), and diffusion models (Stroebel et al., 2023). Types of synthetic realities include:

- Synthetic images – extremely realistic photos of people, events, or places that do not exist (e.g., fake pictures of protests, fabricated evidence), often used in advertising, education, and design, but also for discrediting and manipulating politics.
- Synthetic video (deepfake video) – generated by artificial intelligence, which convincingly changes a person's face, voice, or movements, making them appear to say or do things they never did. Such realistic manipulations of faces and gestures in videos are often used for satirical or propaganda purposes.
- Synthetic audio – voice cloning, AI-generated speech that reproduces a person's unique vocal characteristics, imitating them in a virtually indistinguishable way, allowing the creation of speeches and statements that never took place.
- Virtual persons are completely artificial characters created by artificial intelligence and have a history of existence on the web.
- XR (Extended Reality) environments – including augmented reality (AR), virtual reality (VR), and mixed reality (MR), which combine the physical and digital worlds in real time.

These types of synthetic realities can be combined to create immersive narratives in which every aspect of the story – from text and images to dialogue and voices – is artificially constructed but appears authentic because it appeals to the human tendency to judge credibility based on visual and auditory stimuli.

The best-known and most widespread form of synthetic reality is deepfakes. The term was coined as a combination of the words 'deep learning' and 'fake', which reflects its essence – deep learning in the service of imitation (Odunlami and Banjo, 2025). Deepfake is defined as a product of using machine learning algorithms to create realistic but fake visual materials that can mislead viewers (Das et al., 2025). Most studies noted that deepfake is a form of AI- generated synthetic

media in the form of hyper-realistic images, recordings, and sounds simulating real people (Flores, 2025; Stroebel et al., 2023). Venema and Geradts (2020) focused on the legal and evidentiary aspects, recognising deepfakes as a digital manipulation of audiovisual material that challenges the justice system and the credibility of evidence. Taken together, these definitions present deepfakes as a technology with two faces; on the one hand, an innovative generative tool, and on the other, a potential source of misinformation, a crisis of trust, and threats to the integrity of information. While some applications of deepfakes are harmless – such as entertainment, education, or accessibility applications – their use for political manipulation poses a unique threat to democratic societies.

Deepfakes are exceptionally powerful because the visual signals they present, such as facial expressions and body language, trigger deeply ingrained trust mechanisms, and reinforcement through emotional resonance strengthens memory and belief in the content (Peng et al., 2025). They can therefore engage audiences at both the cognitive and emotional levels, which, combined with a convincing narrative, maximises their persuasive power (Oliullah and Murtuza, 2025). All this makes deepfakes the ideal driver for conspiracy theories, which rely on compelling narratives. By providing visual and audio 'evidence', they make conspiracy theories seem credible, as seeing and hearing something directly strongly influences the formation of beliefs, bypassing traditional scepticism and fact-checking mechanisms. Deepfakes reinforce conspiracy theories in several ways. Conspiracy theories often lack tangible evidence. Deepfakes can deliver this 'evidence', making even the most implausible claims seem authentic (e.g., a fake video showing a politician meeting with foreign secret agents can reinforce an existing narrative of betrayal). At the same time, they influence the rapid spread of conspiracy theories, as visual content is both easy to share and emotionally engaging. As mentioned, social media algorithms also contribute to this by making sensational deepfakes more visible and increasing their reach. On the other hand, the very existence of deepfakes can undermine trust in authentic evidence and allow politically or publicly engaged people to question real recordings and label them as AI-generated. This phenomenon is known as the liar's dividend (Schiff et al., 2025). Thus, synthetic reality can be considered a paradigm shift in the political landscape because it has become a powerful tool for political manipulation and conspiracy mobilisation. By creating convincing evidence, playing with people's senses, and appealing to their emotions, they blur the line between truth and fiction.

Trust and Distrust in Democratic Systems

For almost two decades, we have been witnessing a gradual but steady decline in trust in institutions in Europe (Palacios, 2025; Turska-Kawa, 2025). Trust in public institutions is defined as a cognitive-emotional relationship in which citizens attribute specific competence, intention, and predictability of actions to institutions (Hetherington, 2005). Institutional trust, therefore, presupposes a presumption of goodwill, the ability to act in the public interest, and responsibility for decisions made. A key factor differentiating the level of trust in democratic countries is the perception of the legitimacy of power – not only in a formal and legal sense, but above all in a functional and normative sense. The literature on the subject has repeatedly pointed out that trust in institutions is not a permanent category, but is dynamic in nature, subject to long-term fluctuations, rooted in civic experiences, and sensitive to political disturbances (Affairs and Perry, 2021; Didenko et al., 2020; Marien and Werner, 2019).

Disappointment with the effects of political and economic transformation and ongoing crises – from financial to health-related to military – creates a climate of permanent instability in which the erosion of trust is no longer a random phenomenon, but a systemic one. Multiple crises have contributed to the formation of societies full of uncertainty, fear, and, at the same time, distrust and frustration. All of these problems translate into growing levels of dissatisfaction in society and distrust of institutions. This distrust is fuelled by the poor condition of the entire social system, including the economy, healthcare, military, education, political system, etc. The situation is not improved by political discourse based increasingly on emotions, simplified narratives, manipulation, and even deliberate deconstruction of the concept of truth (Jacob and Milot-Poulin, 2024). In this context, the theory of political agency deserves special attention, in which trust becomes a function of perceived influence on the system. Trust grows if a citizen (or social leader) has a sense of agency. Where there is marginalisation, a lack of dialogue, and exclusion, mistrust arises.

All this, together with the erosion of the separation of powers, increasing polarisation, and the subordination of institutions to political interests, contributes to a deepening crisis of trust. In post-communist countries such as Poland, the issue of trust in institutions is particularly complex. A legacy of distrust of authority, learned political passivity, and deep polarisation of public life are the background and the cause of the observed deficit in the legitimacy of public authority. Therefore, excessive distrust of institutions is a source of disruption in the

functioning of society in a broad sense. If distrust of institutions is widespread, cooperation between individuals becomes complicated and social interactions can be disrupted.

All this pushes people to look for those 'responsible for the whole situation', for answers to why things are so bad. Many find answers in conspiracy theories, which offer a narrative that 'certain hidden groups in power' are to blame. Therefore, recent years, fraught with crises, provide fertile ground for the spread of conspiratorial narratives that offer simple explanations for difficult situations. This provides a sense of security and relief, based on the belief that those who believe in the conspiracy belong to a small group of people who see how the world really works (Adam-Troian et al., 2021; Gligorić et al., 2021).

Conspiracy theories, the dynamic development of information technology, and especially AI-generated content, including deepfakes, pose a serious challenge to trust. Social trust is the glue that holds democracy together, but for that trust to exist, citizens must believe that the information disseminated in public debate is true. When people decide that they cannot believe what they see or hear, a huge problem arises. Fabricated AI-generated content can lead to false beliefs and a feeling that all materials (even videos or audio recordings) may be fabricated. It can even undermine the justification for true beliefs. People will then begin to doubt authentic scientific evidence.

This can have irreversible consequences, including increasing distrust of the news media and delegitimising reliable journalism by associating it with potential falsification. This loss of faith in the media causes people to turn away from the primary source of information and seek information 'on their own', which in many cases pushes them into hermetic information bubbles where they are exposed to conspiracy theories.

What is more, conspiracy theories fuelled by deepfakes can intensify the already significant political and social polarisation. In many US, European, and Polish election campaigns, AI-generated content is used to discredit political opponents or provoke extreme emotions, which mobilise a party's electorate and exacerbate social divisions. Manipulation using deepfakes can influence election results, mobilise the electorate, discredit public figures, or even contribute to geopolitical tensions.

Deepfake technology used as a tool for political manipulation can also make polarised audiences more receptive to fabricated content that slanders the 'other side', while the victims of such attacks reject even justified criticism, calling it 'fake news'. The average citizen who receives conflicting information sinks into 'disbelief in everything', which translates into a lack of social engagement. Such civic apathy can also harm democracy, as it prevents rational debate and those in power from being held accountable. In other words, deepfakes and the disinformation associated with them destroy the common foundation of facts on which healthy public discourse is based and weaken the pillars of democracy – truth and trust – leading to a situation where society remains in permanent doubt. This leads to authentic information being alleged falsifications and rejected. In contrast, falsifications are accepted as probable facts, as long as they fit a particular narrative, causing the erosion of trust in public institutions. This also plays into the hands of the 'enemies of democracy', both internal radicals and external authoritarian regimes, creating ideal conditions for further manipulation and interference.

Materials And Methods

This chapter uses a qualitative research approach, combining case studies with critical content analysis, to understand how synthetic realities support the spread of conspiracy theories and influence social trust in democracies. Selected cases from 2019 to 2025 concerning the use of synthetic media in political contexts – global, European, and Polish – were analysed using media sources, fact-checking reports, and scientific publications. The research process involved identifying synthetic materials, studying their narrative and emotional context, and assessing their socio-political impact. The theoretical part was based on analysing political science, social psychology, and computer science literature. The chosen method ensured reliability and an in-depth understanding of the mechanisms of deepfake use in various political systems.

Results

To better understand how synthetic realities, especially deepfakes, support and perpetuate conspiracy theories, it is worth analysing specific cases in which fake digital materials have become 'evidence' for conspiracy narratives and affect political reality. This mechanism works in two ways: on the one hand, deepfakes can become a tool used to create seemingly credible content that reinforces conspiracy beliefs, but on the other hand, the very awareness of the existence of this technology fosters the emergence of theories about 'replaced' leaders, 'controlled' elites, and 'fake' events. Below are selected examples from recent years, covering various geographical contexts: the world (global cases), the United States, Europe, and Poland. Each case illustrates a different threat aspect – from information warfare, political provocations, and negative election campaigns to local incidents. This demonstrates that the problem of using synthetic realities in politics is multidimensional, and its implications can vary in scope.

Global examples: deepfakes as 'evidence' in theories about false leaders

Gabon (2019): the president's double (digital avatar)
Conspiracy theory: 'The president is dead/no longer in power, and the authorities are hiding this from the public'.
After the publication of a New Year's Eve speech video by the ailing President Ali Bongo, the opposition and part of the public considered it a deepfake. The material was probably authentic. Still, the belief that the government had used artificial intelligence to 'replace' the leader with his digital avatar reinforced the conspiracy narrative about the illegal maintenance of power and led to an attempted coup (Galston, 2020). Consequences of spreading this CT were increased distrust of institutions (in this case, the government), radicalisation of the opposition, and a real risk of destabilisation of the state. The very fact that part of society believed that the video was fake proves that in uncertain times marked by conflict, authoritarian rule, or populism, conspiracy theories about 'digitally generated' leaders can cause real social unrest and violence.

War in Ukraine (2022): Zelensky's capitulation
Conspiracy theory: 'The front-line war recordings are "staged" and all events are fake'.
Early on in Russia's invasion of Ukraine, the first high-profile case of a deepfake was used for propaganda purposes. It was a fake recording of Ukrainian President Volodymyr Zelensky calling for surrender that went viral worldwide, even on hacked Ukrainian television (Wakefield, 2022). The quality of this deepfake was poor (the character had an unnatural voice and stiff motions), so it was quickly exposed and removed from online platforms. Volodymyr Zelensky himself publicly denied the authenticity of the recording. But it still triggered a wave of suspicion about the authenticity of other materials from the front lines, suspecting them of being 'staged'. The consequences of this deepfake were the erosion of trust in information sources and society's susceptibility to narratives about 'Western manipulation'.

Examples from the USA: theories of 'election theft'

Deepfake audio with President Joe Biden (2024)
Conspiracy theory: 'Elections are controlled by technological fraud created by "hidden forces", and citizens' votes do not matter'.
A fake voice message was sent out *en masse* to thousands of voters in New Hampshire before the presidential primaries there. The automated message (robocall) played 'the voice of President Joe Biden' encouraging recipients to 'not vote in the upcoming election' (Matza, 2024). The voice did indeed sound like Biden, so recipients of the message could have been misled. And although the message was illogical, as Biden would be unlikely to discourage his own voters from voting, the mere shadow of doubt is enough to confuse. A mass robocall (reaching nearly 5000 voters) with a voice impersonating the president, discouraging people from voting, was immediately included in narratives about 'manipulation by the electoral elite'. This case showed that a computer-generated voice used to pretend to be a key public figure can be used to try to influence voter turnout and election results. The consequences include fuelling theories of 'stolen elections', increasing polarisation, and legitimising political violence.

Deepfakes during the 2024 presidential campaign in the USA
Due to its intense political polarisation and the importance of social media in public life, the US is particularly vulnerable to the threats posed by deepfakes. And although specific video manipulations targeting politicians had already appeared in previous years (Parkin, 2019.)., e.g., slowing down a recording of Nancy Pelosi to make her look drunk, it was really the 2024 presidential campaign that showed how widely available advanced AI tools for generating fake content would become (Bond, 2024).

Deepfakes and conspiracy theories about Kamala Harris
Conspiracy theory: 'Truth no longer exists, everything is manipulated'.

False claim #1: The Democratic presidential candidate caused an accident that left a 13-year-old girl paralysed and then fled the scene. Information appeared online that Kamala Harris had been accused of causing a car accident in 2011 in which she allegedly injured a teenager and fled the scene. Microsoft specialists discovered that a group called Storm-1516, linked to Russia, created a video, paid an actor to play the role of the alleged victim, and spread the information via a fake website of a non-existent TV station, KBSF-TV, based in San Francisco. The website was created shortly before the first article about the alleged traffic accident was published. The video circulated on social media, including on the X platform (formerly Twitter), with the hashtag #HitAndRunKamala, and was viewed over 2.7 million times.

False claim #2: Kamala Harris is a communist.
The Trump campaign released an AI-generated video showing Kamala Harris and her running mate Tim Walz appearing to pose for a selfie in front of a sign for Revolutionary Communists of America, a far-left group. The goal was not necessarily to create a realistic hoax but to provoke negative emotions among the electorate by associating Harris with communism. This was not an isolated case, as many deepfake memes on this topic were published. According to a DDIA report prepared for the BBC, posts linking Harris to communism gained significant popularity on the Internet after the presidential debate, in which Donald Trump called Kamala Harris and her father 'Marxists', and the term 'Marxist' became popular on social media. Google searches for 'Marxist' in the United States increased by 1,000% in 17 hours. In addition, Trump supporters posted a Russian Communist Party membership card online, allegedly belonging to Kamala Harris. BBC Verify determined that the photo of the membership card came from a website where fake Communist Party documents can be created. This is an example of a

propaganda tool that reinforces negative narratives, even if the audience knows the image is fake. On the other hand, after such a flood of AI-generated memes during the presidential campaign, Donald Trump also fell victim to conspiracy narratives. In response to the confusion surrounding the attempted assassination of Trump in 2024, theories emerged that the attack was a 'false flag', i.e., staged by Trump himself for political gain. The above examples show that in the US, synthetic realities are slowly becoming a common feature of election campaigns, being used both for campaign attacks and as a tool to sabotage the electoral process. Additionally, if an artificial recording can be created, it is also possible to convince one's supporters that inconvenient leaks (e.g., tapes or videos) are the result of intrigues by opponents. Following this narrative, one can question authentic materials and claim that they are fake to protect one's image or undermine the message of the opposing side. The consequences include the disintegration of the public sphere, radicalisation of supporters, polarisation of narratives, and the feeling that 'everything can be fake', leading to real scandals or evidence being ignored.

Europe: deepfakes and narratives of elite betrayal

Slovakia (2023): Šimečka recording

Conspiracy theory: 'Pro-European parties falsify elections and implement foreign interests (implied: Western)'.

During the Slovak parliamentary election campaign, a fake audio file emerged in which the leader of the pro-European party Progresívne Slovensko, Michal Šimečka, allegedly discusses election rigging (e.g., buying votes from Roma) with journalist Monika Tódová (Devine et al., 2024). The recording was published during the so-called election silence period (48 hours before the elections), so despite immediate denials by both sides of the authenticity of the recording, it legally limited the media's ability to respond quickly and refute its content. As a result, despite favourable polls for Šimečka and his party, the Smer party led by Robert Fico, which had a more pro-Russian and anti-Western rhetoric, won. Although there is no hard evidence that this recording 'made the difference' in the elections, it certainly reinforced existing societal narratives about corrupt pro-European politicians and elites' involvement in manipulating election results. It also impacted public trust, as such incidents fuel suspicion of democratic institutions, which deepens political polarisation and erodes confidence in the integrity of elections. In the media and analyses, this case has been cited as one of the first serious tests of the impact of generative technologies (AI/deepfake) on electoral processes in European democracies.

Western Europe (2025): European politicians as puppets of global elites

Conspiracy theory: 'Global elites ruling the world and Europe's subjugation to the United States'.
Disinformation campaigns in Europe increasingly use AI-generated materials (photos and videos) to 'prove' that European political leaders are insignificant and are just puppets in the hands of global elites. Such manipulations, among others, were concerned with Rishi Sunak, Emmanuel Macron, and Olaf Scholz, and were linked to the narrative that European politics is 'rigged' by invisible forces (Łabuz and Nehring, 2024). The most illustrative case occurred on August 18, 2025, when Donald Trump met with Volodymyr Zelensky and European leaders at the White House to discuss the chances of a peaceful end to the war in Ukraine. At that time, an AI-generated photo showing European leaders 'politely sitting' in the White House hallway as if they were humbly waiting for Trump was spread online (Irvine-Brown, 2025). This photo, generated by AI, was intended to ridicule and weaken the image of the EU and reinforce the narrative that European leaders are 'servants' of the US and do not have an independent political position.

Poland: fertile ground for conspiracy theories

Civic Platform spot with the voice of Prime Minister Mateusz Morawiecki (2023)
Conspiracy theory: 'Politicians are lying using new technologies, and voters have no chance of distinguishing truth from fake news'.
In the 2023 Polish parliamentary election campaign, the Civic Platform (Platforma Obywatelska[2]) published a spot using an artificially generated voice of then-Prime Minister Mateusz Morawiecki (Łabuz and Nehring, 2024). The spot featured original fragments of Morawiecki's statements, including those in the Sejm. The 'artificially generated voice' of the prime minister then read excerpts from emails allegedly written by Morawiecki about the difficulties of governing and conflicts with Zbigniew Ziobro, Minister of Justice and leader of the Suwerenna Polska party, a coalition partner of Law and Justice (the ruling party at

2 Polish political party fighting for power, led by Donald Tusk.

the time). It was intended to lend credibility to an ongoing email scandal and undermine the prime minister's credibility. And although the party clearly marked the material as AI-generated content, speculation immediately arose online that it was evidence of media manipulation or electoral fraud. The response to this was to generate 'retaliatory' content, this time with the voice of Donald Tusk, leader of the Civic Platform, intended to undermine the credibility of the opposition leader. The consequences are like those in the US, building the belief that all materials can be manipulated. Deepfake technology in Poland has become a tool of black propaganda, as fake photos of opposition MPs and leaders of opposition parties showing them in discreditable situations (parties, romances, gambling) are also being spread on the Internet, in closed groups and local media. Such 'face pasting' to insinuate scandals reinforces theories about the 'hidden corruption of the elites', which is intended to undermine the reputation of politicians and the parties they represent.

Overall Findings

Based on the case studies and literature review, it can be noted that synthetic realities, i.e., generative technologies based on artificial intelligence (in particular, deepfakes), have significant potential to influence both political processes and the level of public trust in democracies. They have become an integral part of contemporary politics, including election campaigns. Their impact on election results remains limited, but it is significant in social and psychological terms (Łabuz and Nehring, 2024). The use of synthetic realities is multidimensional: on the one hand, it concerns audiovisual manipulation in election campaigns, but on the other, it is part of the spread of conspiracy narratives based on false 'evidence' that exploit the emotions, fears, and prejudices of the audience. Research results show that deepfakes are increasingly used not as direct tools of deception, but as catalysts of distrust, reinforcing the belief that there is no longer any reliable information. The examples presented in the chapter reveal that even content marked as artificially generated triggers a wave of suspicion and accusations of manipulation, contributing to the erosion of standards in public debate. Other examples cited show the use of propaganda memes that perpetuate ideological divisions and reinforce social polarisation despite their obvious fictitiousness.

Another effect of the spread of synthetic media, apart from deceiving audiences, is the loss of the ability to distinguish between truth and falsehood. The consequence is that the authenticity of objective evidence is questioned and labelled as fabricated. So, fabricated material strengthens the so-called 'liar's dividend'. The result is a growing crisis of trust and the susceptibility of societies to manipulation. With this, deepfakes not only support existing conspiracy theories but also become the basis for new narratives themselves.

Encouragingly, synthetic reality technologies have not caused an 'information apocalypse'. Still, they have contributed to a systemic shift in the balance between trust and suspicion in the public domain, including the political sphere. The key challenge is not the battle against the technology itself, legal bans, or the development of detection tools, but the rebuilding of societies' epistemic resilience through media education, source transparency, and strengthening citizens' digital skills. After all, the use of synthetic realities to support conspiracy narratives can be seen as a potential erosion of the foundations of social trust in democratic processes.

References

Babaei, R., S. Cheng, R. Duan and S. Zhao. (2025). "Generative Artificial Intelligence and the Evolving Challenge of Deepfake Detection: A Systematic Analysis." *Journal of Sensor and Actuator Networks* 14 (1): 17.

Boncu, Ș., O. Prundeanu, A. C. Holman and S. A. Popușoi. (2022). "Believing in or Denying Climate Change for Questionable Reasons: Generic Conspiracist Beliefs, Personality, and Climate Change Perceptions of Romanian University Students." *International Journal of Environmental Research and Public Health* 19 (24): 17038.

Bond, S. (2024). "How AI deepfakes polluted elections in 2024." *NPR.* Available at https://www.npr.org/2024/12/21/nx-s1-5220301/deepfakes-memes-artificial-intelligence-elections.

Cardenuto, J. P., J. Yang, R. Padilha, R. Wan, D. Moreira, H. Li, S. Wang, F. Andaló, S. Marcel and A. Rocha. (2023). "The Age of Synthetic Realities: Challenges and Opportunities." *APSIPA Transactions on Signal and Information Processing* 12 (1).

Das, S. S., M. Agarwal, S. Rajan and A. Padhi. (2025). "Synthetic realities: Youth media literacy and trust in the age of digital deception." *Information Development,* available at https://journals.sagepub.com/doi/abs/10.1177/02666669251374658.

Devine, C., D. O'Sullivan and S. Lyngaas. (2024). "A fake recording of a candidate saying he'd rigged the election went viral. Experts say it's only the beginning." *CNN Politics.* Available at https://edition.cnn.com/2024/02/01/politics/election-deepfake-threats-invs.

Didenko, N. I., G. F. Romashkina, D. F. Skripnuk and S. V. Kulik. (2020). "Dynamics of Trust in Institutions, the Legitimacy of the Social Order, and Social Open Innovation." *Journal of Open Innovation: Technology, Market, and Complexity* 6 (4): 111.

Douglas, K. M., J. E. Uscinski, R. M. Sutton, A. Cichocka, T. Nefes, C. S. Ang and F. Deravi. (2019). "Understanding Conspiracy Theories." *Political Psychology* 40 (S1): 3–35.

Filho, L. L. and A. Rocha. (2024). "The Rise of Synthetic Realities: Impact, Advancements, and Ethical Considerations." *IEEE SMC Magazine.* Available at https://www.ieeesmc.org/wp-content/uploads/2024/03/FeatureArticle_03_2024.pdf.

Flores, A. (2025). "Synthetic Realities: How Deepfakes Challenge Consent in the Digital Age." *Social Science Research Network* (SSRN paper no. 5368796).

Galston, W. (2020). "Is seeing still believing? The deepfake challenge to truth in politics." *Brookings.* Available at https://www.brookings.edu/articles/is-seeing-still-believing-the-deepfake-challenge-to-truth-in-politics/.

George, D. A. S. and A. S. H. George. (2023). "Deepfakes: The Evolution of Hyperrealistic Media Manipulation." *Partners Universal Innovative Research Publication (PUIRP)* 1 (2): 58–74.

Hetherington, M. J. (2005). *Why Trust Matters: Declining Political Trust and the Demise of American Liberalism.* Princeton: Princeton University Press.

Hynek, N., B. Gavurova and M. Kubak. (2025). "Risks and benefits of artificial intelligence deepfakes: Systematic review and comparison of public attitudes in seven European countries." *Journal of Innovation and Knowledge* 10 (5): 100782.

Irvine-Brown, R. (2025). "How a fake image went round the world in a few hours". *BBC News.* Available at https://www.bbc.com/news/live/cx29wjvjnp2t?post=asset%3A0b621af4-f5d9-4e71-bd92-c20d6752815e#post.

Jacob, S. and J. Milot-Poulin. (2024). "Lies and politics: Until death do us part..." In *Evaluation in the Post-Truth World*, eds. Marra, M., K. Olejniczak and A. Paulson, 57–75. London: Routledge.

Łabuz, M. and C. Nehring. (2024). "On the way to deep fake democracy? Deep fakes in election campaigns in 2023." *European Political Science* 23 (4): 454–473.

Marien, S. and H. Werner. (2019). "Fair treatment, fair play? The relationship between fair treatment perceptions, political trust and compliant and cooperative attitudes cross-nationally." *European Journal of Political Research* 58 (1): 72–95.

Matza, M. (2024). "Fake Biden robocall tells voters to skip New Hampshire primary election." *BBC News.* Available at https://www.bbc.com/news/world-us-canada-68064247.

Odunlami, O. and O. A. Banjo. (2025). "Deepfakes and the Crisis of Trust: Public Perception of Media Authenticity in the Age of Synthetic Content." *Nigerian Journal of Technical Education* 24 (2): 45–59.

Oliullah, M. and H. M. Murtuza. (2025). "Exploring the Impact of Deepfakes on Cognitive Processing in Political Contexts: A Systematic Review." *Social Science Research Network,* SSRN Paper no. 5390735.

Palacios, I. (2025). "How democratic backsliding and populism affect trust in democratic institutions." *Democratization* (online): 1–24.

Parkin, S. (2019). "The rise of the deepfake and the threat to democracy." *The Guardian,* 4 October 2025. Available at http://www.theguardian.com/technology/ng-interactive/2019/jun/22/the-rise-of-the-deepfake-and-the-threat-to-democracy.

Peng, Q., Y. Lu, Y. Peng, S. Qian, X. Liu and C. Shen. (2025). "Crafting Synthetic Realities: Examining Visual Realism and Misinformation Potential of Photorealistic AI-Generated Images." *Proceedings of the Extended Abstracts of the CHI Conference on Human Factors in Computing Systems*: 1–12.

Pilch, I., A. Turska-Kawa, P. Wardawy, A. Olszanecka-Marmola and W. Smołkowska-Jędo. (2023). "Contemporary trends in psychological research on conspiracy beliefs: A systematic review." *Frontiers in Psychology* 14: 1075779.

Schiff, K. J., D. S. Schiff and N. S. Bueno. (2025). "The Liar's Dividend: Can Politicians Claim Misinformation to Evade Accountability?" *American Political Science Review* 119 (1): 71–90.

Sharma, V. K., R. Garg and Q. Caudron. (2025). "A systematic literature review on deepfake detection techniques." *Multimedia Tools and Applications* 84 (20): 22187–22229.

Stroebel, L., M. Llewellyn, T. Hartley, T. S. Ip and M. Ahmed. (2023). "A systematic literature review on the effectiveness of deepfake detection techniques." *Journal of Cyber Security Technology* 7 (2): 83–113.

Sunstein, C. R. and A. Vermeule. (2009). "Conspiracy theories: Causes and cures." *The Journal of Political Philosophy* 17 (2): 202–227.

Swami, V. and A. Furnham. (2014). "Political paranoia and conspiracy theories." In *Power, Politics, and Paranoia: Why People Are Suspicious About Their Leaders,* eds. van Prooijen, J. W. and P. A. M. van Lange, 218–236. Cambridge: Cambridge University Press.

Turska-Kawa, A. (2025). *Podmiotowość polityczna w erze depolityzacji.* Katowice: Wydawnictwo Uniwersytetu Śląskiego.

United Nations Department of Economic and Social Affairs and Perry, J. (2021). "Trust in Public Institutions: Trends and Implications for Economic Security." *United Nations Department of Economic and Social Affairs Policy Briefs.*

Uscinski, J. E., C. Klofstad and M. D. Atkinson. (2016). "What Drives Conspiratorial Beliefs? The Role of Informational Cues and Predispositions." *Political Research Quarterly* 69 (1): 57–71.

Van Prooijen, J.-W. and K. M. Douglas. (2017). "Conspiracy theories as part of history: The role of societal crisis situations." *Memory Studies* 10 (3): 323–333.

Venema, A. E. and Z. J. Geradts. (2020). "Digital Forensics, Deepfakes and the Legal Process." *SciTech Lawyer* 16 (4): 14–23.

Wakefield, J. (2022). "Deepfake presidents used in Russia-Ukraine war." *BBC News.* Available at https://www.bbc.com/news/technology-60780142.

Wang, Y. (2023). "Synthetic Realities in the Digital Age: Navigating the Opportunities and Challenges of AI-Generated Content." *Institute of Electrical and Electronics Engineers (IEEE).*

Zonis, M. and C. M. Joseph. (1994). "Conspiracy thinking in the Middle East." *Political Psychology* 15 (3): 443–459.